

Hazard/Risk Assessment

TWAc-Check: A New Approach to Determine the Appropriate Use of Time-Weighted Average Concentration in Aquatic Risk Assessment

Yvonne Wolf,^{a,*} Faten Gabsi,^{b,c} Eric Bruns,^a Simon Heine,^a Andreas Solga,^a Johannes Witt,^a and Thomas G. Preuss^a^aBayer, Monheim, Germany^bRifcon, Hirschberg an der Bergstrasse, Germany^cRegachem Consulting, Sfax, Tunisia

Abstract: In pesticide risk assessment, regulatory acceptable concentrations for surface water bodies (RAC_{sw,ch}) are used that are derived from standard studies with continuous exposure of organisms to a test compound for days or months. These RAC_{sw,ch} are compared with the maximum tested concentration of more realistic exposure scenarios. However, the actual exposure duration could be notably shorter (e.g., hours) than the standard study, which intentionally leads to an overly conservative Tier 1 risk assessment. This discrepancy can be addressed in a risk assessment using the time-weighted average concentration (TWAc). In Europe, the applicability of TWAc for a particular risk assessment is evaluated using a complex decision scheme, which has been controversial; thus we propose an alternative approach: We used TWAc-check (which is based on the idea that the TWAc concept is just a model for aquatic risk assessment) to test whether the use of a TWAc is appropriate for such assessment. The TWAc-check method works by using predicted–measured diagrams to test how well the TWAc model predicts experimental data from peak exposure experiments. Overestimated effects are accepted because the conservatism of the TWAc model is prioritized over the goodness of fit. We illustrate the applicability of TWAc-check by applying it to various data sets for different species and substances. We demonstrate that the applicability is case dependent. Specifically, TWAc-check correctly identifies that the use of TWAc is not appropriate for early onset of effects or delayed effects. The proposed concept shows that the time window is a decisive factor as to whether or not the model is acceptable and that this concept can be used as a potential refinement option prior to the use of toxicokinetic-toxicodynamic models. *Environ Toxicol Chem* 2022;41:1778–1787. © 2022 Bayer AG. *Environmental Toxicology and Chemistry* published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Risk assessment; Pesticide risk assessment; Pesticide regulation; Time-weighted average; European Food Safety Authority aquatic guidance document; Conservatism

INTRODUCTION

In standard studies aiming to derive regulatorily acceptable concentrations for surface water bodies (RAC_{sw,ch}) for pesticide risk assessment, organisms are constantly exposed to a test compound over several days, weeks, or even months.

According to the European Food Safety Authority Aquatic Guidance Document (EFSA AGD; EFSA, 2013), in Tier 1 risk assessment the RAC_{sw,ch} values derived from this long-term exposure are compared with the maximum concentrations of more complex exposure scenarios (predicted environmental concentration [PEC]_{sw,max}). Depending on the defined product, good agricultural practice, exposure scenario, and characteristics of the compound, the exposure event that defines the PEC_{sw,max} can have a duration notably shorter than the exposure period in the effect study, for example, only a few hours. This situation intentionally leads to an overly conservative Tier 1 risk assessment. A possible technique to address this discrepancy is to conduct a risk assessment based on time-weighted average concentration (TWAc). The TWAc is calculated by determining the area under the curve

This article includes online-only Supporting Information.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Yvonne Wolf and Faten Gabsi contributed equally to this study.

* Address correspondence to: yvonne.wolf@bayer.com

Published online 18 April 2022 in Wiley Online Library

(wileyonlinelibrary.com).

DOI: 10.1002/etc.5346

concentration within the chosen time window and dividing the results by the window length (i.e., 7 days; EFSA, 2013). The TWAc method has been applied in various situations other than pesticide risk assessment. These include general assessment of the toxic profiles of compounds under time-variable exposure concentrations (Belgers et al., 2011; Watanabe et al., 2018; Zimmer et al., 2018) and assessment of fluctuating concentrations of different compounds in natural habitats by passive samplers for comparison with laboratory experiments, for example (Valenzuela et al., 2021).

Within the European Union Workshop on Linking Aquatic Exposure and Effects in the Registration Procedure of Plant Protection Products (ELINK; Brock et al., 2010), which aimed to improve guidance on linking exposure and effects in the aquatic risk assessment for pesticides in Europe, it was clarified that under specific conditions (for example, delayed effects), the TWAc should not be used because it is not sufficiently conservative. A complex decision scheme is given in the EFSA AGD to evaluate the applicability of the TWAc approach (EFSA, 2013). Recently, a number of procedural questions around the demonstration of the use of the TWAc with regard to a specific substance have been raised in complex and in part controversial discussions on how to practically handle certain parts of the EFSA AGD decision scheme for the TWAc (EFSA, 2013, 2015). A final clarification on these matters is not yet available.

Therefore, we propose an alternative approach, TWAc-check, to determine whether the use of TWAc is appropriate in aquatic risk assessment. The TWAc-check method is based on the idea that the TWAc concept is just a model for aquatic risk assessment in that it uses a simplifying assumption to allow the prediction of effects under time-variable exposure. Specifically, the general TWAc approach assumes that the effect seen with a time-variable exposure is the same as for a constant exposure at the average concentration. Whether this assumption can be made is most likely species and substance specific. The decision scheme applied so far (EFSA, 2013) investigates certain features (e.g., linear reciprocity, delayed effects, early onset of effects) to decide whether the use of the TWAc is appropriate. So far, there is no agreed methodology on how these individual features should be evaluated.

In contrast, the TWAc-check we suggest is straightforward, using the underlying predictions of the general TWAc approach to test whether the TWAc is conservative/realistic for a specific

species and substance. The approach is illustrated using case studies of different species, macrophyte (*Lemna gibba*), invertebrate (*Daphnia magna*), and vertebrate (*Danio rerio*), exposed to chemicals with different mechanisms of action.

MATERIALS AND METHODS

Description of TWAc-check

Figure 1 provides an illustration of the proposed TWAc-check method: Two different experiments are needed, a standard exposure experiment on the species driving the risk assessment and at least one refined (pulse) exposure test. From the standard study, the concentration–response curve is generated, and its properties (median effective concentration [EC50] and slope) are determined. In our examples we used the probit model (Bliss, 1934) and the ToxratPro software (Ver 3.2.1). However, other types of concentration–response curves can also be used. For each exposure concentration of the pulse exposure test, the TWAc is derived using a time window of 7 days. This is the default value recommended by ELINK (Brock et al., 2010) and adopted in the EFSA AGD (2013) for invertebrates, fish, and macrophytes. However, as stated in the EFSA AGD (2013), the length of the time window can be reduced to increase the conservatism of the model, or can be prolonged for the opposite effect. This TWAc is used in the concentration–response function derived from the chronic study with constant exposure to predict the effect for the short-term (pulse) exposure event. The predicted effects using the TWAc of the pulse exposure study and the concentration–response curve parameter from the concentration–response relationship of the constant exposure test are plotted against the measured effects at the respective concentrations in the pulse exposure test at day x , with x corresponding to duration of the chronic experiment (i.e., 7 days for *Lemna*, 21 days for *Daphnia*, and 28 days for fish). This direct comparison via TWAc-check evaluates whether the TWAc model is realistic, conservative, or not acceptable for predicting effects of short-term exposure events based on constant long-term exposure data. To account for the variability of the data, the 95% confidence intervals are plotted alongside the data. Depending on the type of data, that is, quantal or graded (Ashauer et al., 2011), the 95% confidence interval should be calculated with the corresponding statistical method.

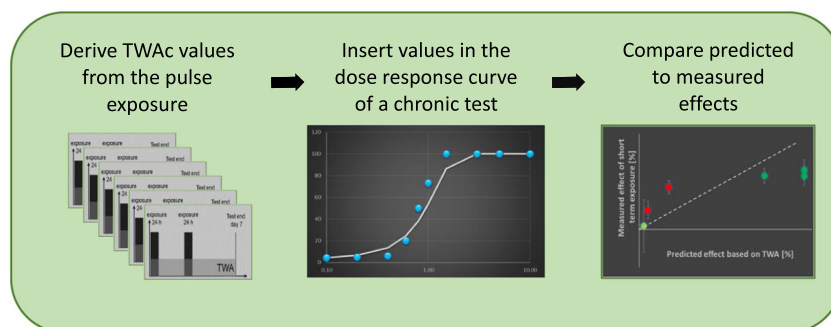


FIGURE 1: Example illustrating the stepwise time-weighted average (TWA) approach. TWAc, time-weighted average concentration.

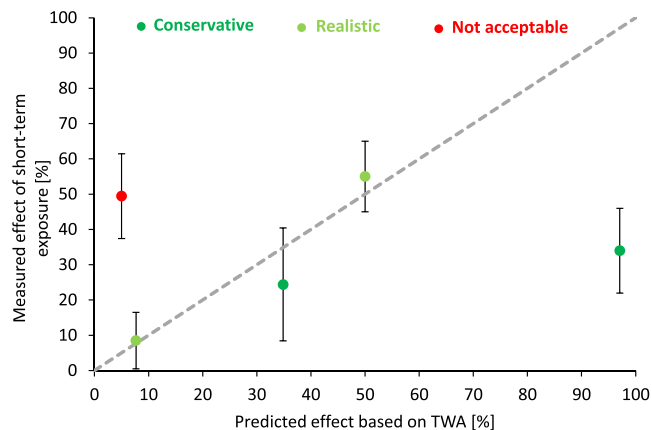


FIGURE 2: Example of predicted effects based on time-weighted average (TWA) concentration versus measured effects of a short-term exposure event.

The predicted–measured diagrams are then assessed in the following way: Data points and confidence intervals lying strictly in the upper part of the 1:1 line (shown in red in Figure 2) indicate that the TWAc underestimates effects. Predictions lying strictly below the 1:1 line (in dark green) mean that the model overestimates effects and thus the TWAc is conservative. Data points lying above or below the 1:1 line, but with confidence intervals crossing that line, are proposed to be realistic and are shown in light green. Therefore, data points in dark green and light green indicate acceptable predictions by TWAc-check and an appropriate use of the TWAc in risk assessment. In this way, we prioritize the conservativeness of model predictions rather than their accuracy. We accept a threshold tolerance of only 10% of unacceptable predictions (strictly above the 1:1 line). In the Supporting Information, we provide the Excel spreadsheet describing the stepwise approach including calculations for the seven case studies described in the present study as well as a template Excel spreadsheet for applying TWAc-check.

To summarize, TWAc-check consists of the following steps:

1. Plot the predicted–measured diagram: For each concentration level in the pulse exposure study:
 - a. Calculate the 7-day TWAc from the pulse exposure experiment(s);
 - b. Insert the TWAc into the standard concentration–response relationship to obtain the predicted effect;
 - c. Plot the predicted versus the observed effect into the predicted–measured diagrams together with the 95% confidence intervals of the measured data.
2. Determine the percentage of unacceptable predictions (i.e., points in the predicted–measured diagram with 95% confidence interval strictly above the 1:1 line). A TWAc is considered acceptable for a specific substance and species if there are no more than 10% unacceptable predictions.

Potential modifications to this procedure are discussed in the *Results and Discussion* section.

Test of TWAc-check on theoretical data showing early onset and delayed effects

Before the use of TWAc-check, it is critical to test whether the method correctly identifies situations in which the TWAc model is not applicable (EFSA, 2013). Within the EFSA AGD, two scenarios are specified for which TWAc would not be applicable, early onset of effects (Figure 3A) and delayed effects (Figure 3B).

We created two artificial examples for these extreme scenarios, with a constant exposure test and a pulse exposure test both lasting 14 days. For simulation of early onset of effects in chronic experiments, effects appear at day 2 in the chronic study and after that stay constant. In the pulse exposure study the effects would also occur at day 2 and stay constant afterward. So the results of both experiments are similar.

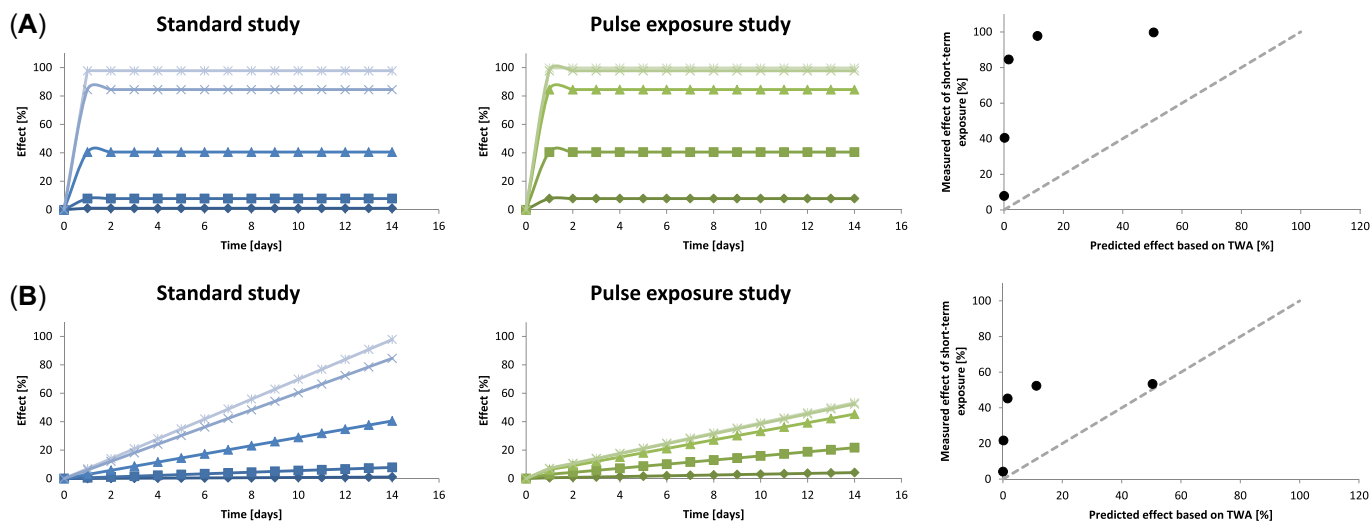


FIGURE 3: Validation of the use of time-weighted average (TWA) concentration in theoretical extreme examples for (A) early onset of effects and (B) latency of effects.

For delayed effects, effects in the pulse exposure studies would last longer than the exposure. This was simulated using a constant increase in the observed effect in both the constant and the pulse exposure studies. However, effects in the pulse exposure scenario increase less than in the constant exposure scenario.

Test of TWAc-check on experimental data

We tested TWAc-check in different scenarios including the effects of different herbicides on *L. gibba* (foramsulfuron, iodosulfuron-methyl-sodium, mesosulfuron-methyl, and thien carbazon-methyl) and *D. magna* (phenmedipham and aclonifen) in addition to effects of the fungicide spiroxamine on *D. rerio*. The two exposure scenarios were constant exposure, and 1–3-day peak exposure.

Inhibition of growth on *L. gibba* was investigated (using test guideline 221 of the Organisation for Economic Co-operation and Development [OECD], 2002). Macrophyte studies revealed that the most sensitive endpoint was frond area for mesosulfuron-methyl and frond number for the other compounds. A short description of these studies is provided in Table 1. Calculation files with experimental data used in each case study are provided in the Supporting Information. The concentration–response curves for the response variable “7-day growth rate” and basic parameters (EC50 and slope; see Table 1) as predicted by the probit regression model were derived and used to predict effects from the pulse exposure tests.

In addition, TWAc-check was used to test phenmedipham and aclonifen for their potential to inhibit reproduction (brood size) in *D. magna* (test guideline 211; OECD, 2012). A short description of these studies is provided in Table 2. Calculation files are provided in the Supporting Information. Three studies were considered for TWAc-check of aclonifen in *D. magna*. The inhibition of reproductive output was recorded in a 21-day reproduction study; the concentration–response parameters were estimated based on these results, and then the effects of the pulse exposure studies were predicted.

Two fish full life cycle studies (Table 3) of juvenile *D. rerio* exposed to spiroxamine (Bayer, 2009, 2014b) were used to test the applicability of TWAc-check. Fish survival was the most sensitive endpoint in both studies. Concentration–response parameters were derived from a constant exposure study (Bayer, 2009) for survival of juvenile fish after 28 days and were used to predict the effects from a peak exposure study (Bayer, 2014b). A 7-day TWA was chosen for the assessment of the TWA approach.

The full study reports can be requested for noncommercial use via the Bayer CropScience transparency initiative (email: cropscience-transparency@bayer.com) by referencing the present study and citing the respective M-Number of the study by Bayer (2009). This registration step is necessary to prevent commercial use of the studies by competitors of the study owner.

RESULTS AND DISCUSSION

The present study proposes a new approach, TWAc-check, to test the protectiveness of TWAc in aquatic risk assessment of

pesticides using results from refined exposure experiments. Overall, TWAc-check is not data hungry and can be tested using one chronic experiment, which is generally performed to support an aquatic risk assessment, and one pulse exposure test, often conducted as a Tier 2c refinement option. This means that we can extract valuable information from and make further use of the already existing data. Furthermore, TWAc-check is easy to apply and reproduce. Proof that the TWAc is fit for purpose for a specific compound and species can be shown by comparing the predicted effects at the TWAc using the concentration–response properties from a standard test type with the measured effects from an additional pulse exposure experiment. This approach complies with recommendations to validate a method in both a compound- and a species-specific way. A successful validation shows that the TWAc is able to realistically or conservatively predict effects at time-variable exposure from constant exposure. The TWAc-check uses predicted–measured diagrams to decide whether the TWAc yields regulatory acceptable (i.e., realistic or conservative) predictions for the specific substance and species at hand. Specifically, this means that data points that are clearly below the 1:1 line overestimate the measured effect and thus can be considered regulatorily conservative. In the present study, we propose to accept no more than 10% nonprotective estimations to ensure that TWAc is conservative.

Test TWAc-check on early onset of effects and delayed effects

As an initial plausibility check, we applied TWAc-check to hypothetical effects from data representing early onset of effects (Figure 3A) and delayed effects (Figure 3B). The time to onset of effects is investigated (for macrophytes, for example) to reveal how rapidly a compound affects the plants. This is done to exclude the possibility that short-term exposure to high concentrations such as may occur under realistic outdoor conditions (e.g., drift or runoff events) produces effects that are “overlooked” when comparing the Tier 1 RAC_{sw,ch} values with the averaged (and thus lowered) TWAc. Investigating “non”-latency of effects is performed to prove that delayed effects in the postexposure phase caused by damage during exposure are not to be expected. Therefore, any approach for testing the appropriateness of TWAc needs to identify TWAc as not suitable in the scenarios with latency of effects or early onset of effects. In our hypothetical examples, the measured effects in the pulse exposure experiment were always higher than those calculated using the TWAc. In fact, all points lay above the 1:1 line. Thus, TWAc-check identifies that the application of the TWAc is not protective for both scenarios, because all predictions were underestimating the effects (Figure 3). This is in accordance with the EFSA (2013) statement that for these two extreme cases, using the TWAc is not appropriate. The TWAc-check proved its ability to detect scenarios in which the TWAc cannot be applied, and therefore we can confidently use it on real data sets.

As a second step, we used TWAc-check on real example applications for which the PEC_{sw} exceeded the RAC, implying that an exposure refinement was necessary.

TABLE 1: Studies used in the modeling for macrophytes

Compound	Experiment type	Study duration	Exposure duration	Tested concentrations	Endpoint	Reference	EC50	Slope
Foramsulfuron	Semistatic study	6 weeks ^a	7 days	0.2, 0.4, 0.8, 1.6, and 3.2 µg a.s./L (nominal)	Frond number at day 7 (growth rate)	Bayer (2013a)	1.22	-1.29
	1 + 6-day pulse exposure test	7 days	1 day	0.50, 1.10, 2.42, 5.32, 11.70, 25.80, and 56.70 µg a.s./L (nominal)		Bayer (2013b)	—	—
	2 + 5-day pulse exposure test	7 days	2 days	1.30, 3.24, 8.06, 20.10, and 50.00 µg a.s./L (nominal)		Bayer (2017; Design 1)	—	—
Thiocarbazon-methyl	Semistatic study	7 days	7 days	0.086, 0.209, 0.542, 1.26, 3.06, and 7.7 µg a.s./L (measured)	Frond number at day 7 (growth rate)	Bayer (2006)	1.36	-2.736
	1 + 6-day pulse exposure test	7 days	1 day	2.21, 4.90, 10.8, 25.00, and 56.14 µg a.s./L (measured)		Bayer (2016a; Design 2)	—	—
	2 + 5-day pulse exposure test	7 days	2 days	2.13, 4.66, 10.30, 24.18, and 53.80 µg a.s./L (measured)		Bayer (2016a; Design 1)	—	—
Iodosulfuron-methyl-sodium	Semistatic study	6 weeks ^a	7 days	0.1, 0.2, 0.4, 0.8, and 1.6 µg a.s./L (measured)	Frond number at day 7 (growth rate)	Bayer (2013c)	1.08	-3.35
	2 + 5-day pulse exposure test	7 days	2 × 1 day	1.91, 6.1, 19.5, 62.5, and 200 µg a.s./L (measured)		Bayer (2016b; Design 1)	—	—
	1 + 6-day pulse exposure test	7 days	1 day			Bayer (2016b; Design 2)	—	—
Mesosulfuron-methyl	Semistatic study	8 weeks ^a	7 days	0.194, 0.388, 0.775, 1.55, and 3.1 µg a.s./L (measured)	Frond area at day 7 (growth rate)	Bayer (2013d)	1.29	-2.53
	2 + 5-day pulsed exposure test	7 days	2 × 1 day	1.23, 3.7, 11.1, 33.3, 100, 25.8, 56.7 µg a.s./L (measured)		Bayer (2016c; Design 1)	—	—
	1 + 6-day pulse exposure test	7 days	1 day	1.23, 3.7, 11.1, 33.3, 100, 25.8, 56.7 µg a.s./L (measured)		Bayer (2016c; Design 2)	—	—

Design 1: 24-h peaks on day 0 and day 3, and thus 48-h exposure in total; study duration 7 days. Design 2: 24-h peaks on day 0 and day 7, only first week accounted for in the present study, which is then equivalent to a 24-h (single-peak) exposure study with 7 days total duration; overall study duration 14 days.

^aIn chronic bioassays, only the first week was used in the modeling.

EC50 = median effective concentration; a.s. = active substance.

TABLE 2: Studies used in the modeling for daphnids

Compound	Experiment type	Study duration	Exposure duration	Tested concentrations (measured)	Endpoint	Reference	EC50	Slope
Phenmedipham	Chronic toxicity under flow-through conditions Pulse exposure	21 days	21 days	5, 10.5, 21.6, 45.8, and 78.4 µg a.s./L	Inhibition of reproduction	Bayer (2014a)	48.8	-3.7
		21 days	3 days (study Days 0 to 3)	22.1, 22.2, 41.2, 43.2, 85.2, and 87.5 µg a.s./L				
Aclonifen	Semistatic chronic toxicity study 2-day pulse exposure study 2 + 1-day pulse exposure study	21 days	21 days	6.27, 14.2, 34.8, 78.1, and 212 µg a.s./L	Inhibition of reproduction	Bayer (2016e)	47.5	-4.5
		21 days	2 days ^a	241, 393, 566, 855, and 1324 µg a.s./L				
		21 days	3 days ^b	213, 335, 516, 807, and 1248 µg a.s./L				

^aDesign 2: 48-h peaks on day 0 and day 14, only first week used, thus 48-h exposure in total.
^bDesign 3: 48-h peaks on day 0 and day 7, only first week used, and thus 72-h exposure in total.
 EC50 = median effective concentration.

Test of TWAc-check on experimental data

Effects of foramsulfuron on growth rate (frond number) directly observed in both refined exposure experiments (1- and 2-day pulse exposures; Figure 4A) were either realistic or conservative, because all model predictions with their confidence levels were below or cross the 1:1 line. For iodosulfuron-methyl-sodium, the effects observed on growth rate (frond number; Figure 4B) in the pulse exposure experiments were lower than those predicted by the TWAc model using a 7-day time window at both the low and high exposure concentrations. The model only underestimated the effect at the lowest tested concentration of the 2-day pulse exposure scenario (1.91 µg/L). However, the confidence limit hit the 1:1 line and is thus considered to be a realistic prediction. Results for thien carbazole-methyl showed that predicted inhibition of growth rate (frond number; Figure 4C) by the 7-day TWAc-check at the lowest concentrations (2.21 and 2.13 µg/L for 1- and 2-day pulse scenarios, respectively) and their confidence levels were entirely located above the 1:1 line, and thus the measured effects were underestimated. All other data points were either realistic or conservative. At the higher concentrations, growth rate inhibition was overestimated by TWAc-check. Because the TWAc underestimated effects at the lowest concentration and in both scenarios (representing 20% of the entire data set), nonprotective use of the TWAc was demonstrated for this compound and *L. gibba*. For mesosulfuron-methyl (Figure 4D), TWAc-check was able to predict the effects of pulse exposure realistically at the lowest concentration (1.23 µg/L), and overestimated the effects at concentrations higher than 3.7 µg/L for both 1- and 2-day pulse exposure scenarios.

With respect to the 10% threshold rule of our assessment, and based on the 7-day TWAc, the use of TWAc was justified for effects on growth rate of *L. gibba* with all compounds, except for thien carbazole-methyl, for which the lowest concentrations failed (Figure 4).

Phenmedipham effects on *D. magna* directly observed in the pulse exposure experiment were mainly smaller than those predicted from the constant exposure study (data points right of the dashed line; Figure 5A), thus indicating that the TWAc was realistic or conservative. Two data points (in light green) were above the 1:1 line, showing large confidence intervals; however, their range reached the acceptable area. A comparison of the predicted and measured effects of aclonifen on *D. magna* based on two independent pulse experiments each including two separate pulses (Figure 5B) shows that all data points were below the 1:1 line; that is, the measured effects in the pulse experiments were always lower than those predicted based on a constant exposure to aclonifen.

These results show that reproductive effects on daphnids (reduction in the number of offspring) for both phenmedipham and aclonifen were similar whether they were exposed for a short time to a high concentration or for a longer time to a lower concentration. This justifies the use of the 7-day TWAc for *D. magna* exposed to each of these compounds.

For fish, the most sensitive endpoint was survival of the F1 generation. The highest effect was obtained after exposure of

TABLE 3: Studies used in the modeling for fish

Compound	Experiment type	Study duration	Exposure duration	Tested concentrations (measured)	Endpoint	Reference	EC50	Slope
Spiroxamine	Chronic toxicity under flow-through conditions	35 days	35 days	0.1, 2.6, 6.4, and 16 µg a.s./L	Survival of juveniles	Bayer (2009)	19.9	-1.278
	2-day pulse exposure study	28 days	2-day (study days 0–14)	15.8, 30.4, 63.9, and 255 µg a.s./L		Bayer (2014b)	—	—

EC50 = median effective concentration.

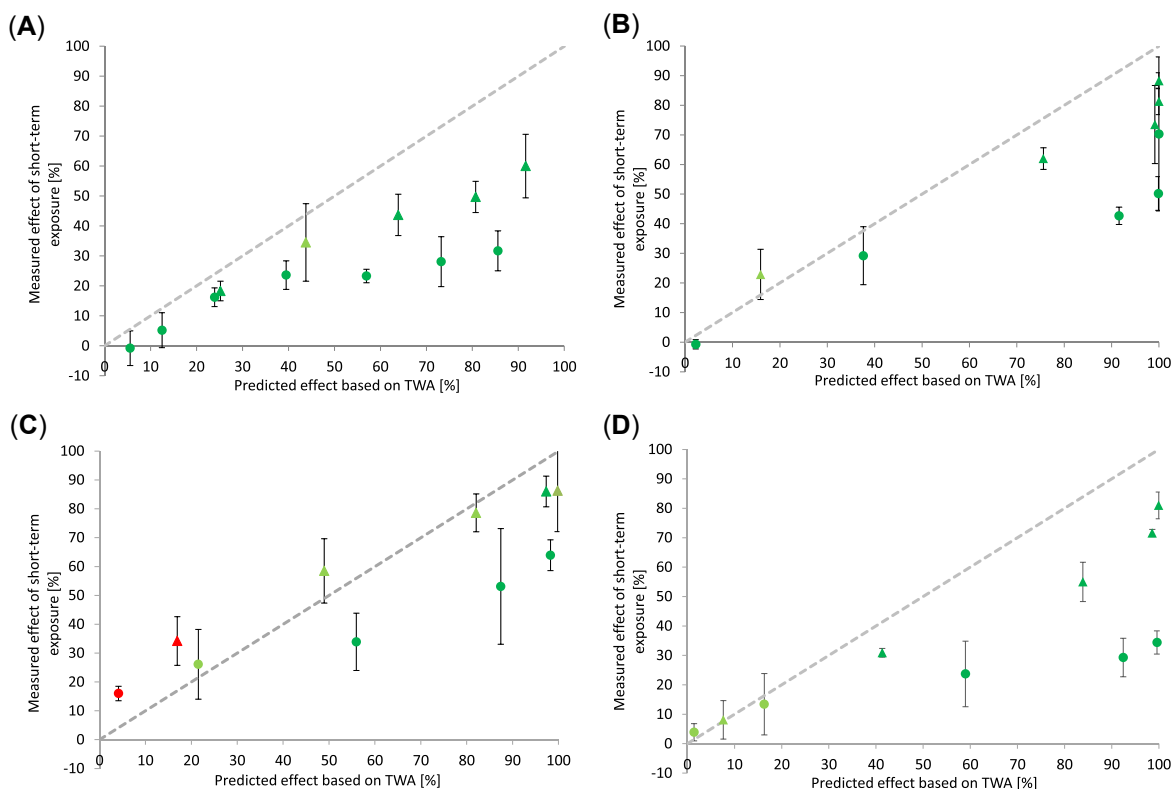


FIGURE 4: Predicted–measured diagrams of observed inhibitions (frond number/area; growth rate) of *Lemna gibba* for foramsulfuron (A), iodosulfuron-methyl-sodium (B), thiocarbazone-methyl (C), and mesosulfuron-methyl (D) based on a 7-day time-weighted average (TWA) concentration; triangles and dots indicate data from different test designs.

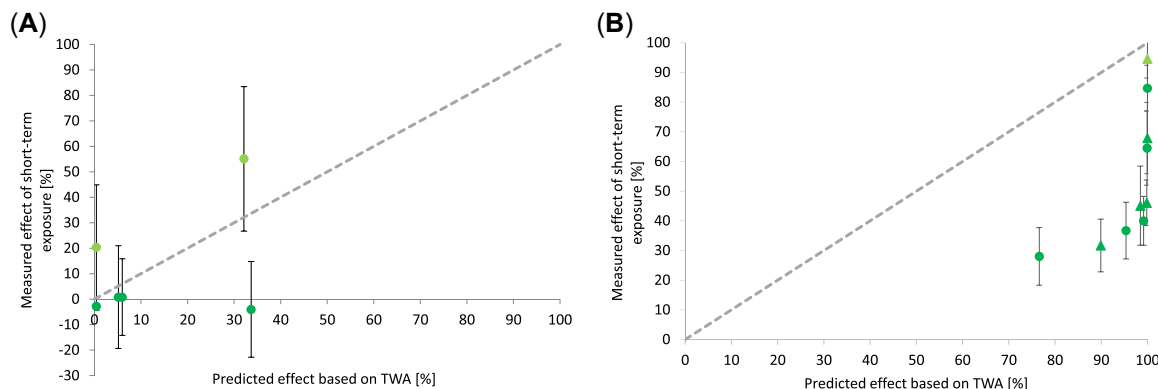


FIGURE 5: Predicted–measured diagrams of total living offspring/living *Daphnia magna* for phenmedipham (A) and aclonifen (B) based on a 7-day time-weighted average (TWA) concentration; triangles and dots indicate data from different test designs.

adult fish in their reproductive phase. Figure 6 shows that effects on juvenile survival observed in the pulse exposure experiment for spiroxamine were lower than would be predicted using the toxicity parameters of the constant exposure study at low exposure concentrations. At higher test concentrations, the TWAc model slightly underestimated the effects in the pulse exposure experiment, but the range of the confidence levels reached the 1:1 line. Overall, the use of TWAc did not underestimate the level of toxicity. The TWAc-check justified the use of the TWAc for a short-term exposure event in a water–sediment system with a 7-day TWAc, because the confidence interval of all data points crossed the 1:1 line.

Prediction results using simulations of hypothetical effects and those from real case studies together indicate that TWAc-check is able to distinguish between acceptable and non-acceptable uses of the TWAc. We believe that our proposed threshold of 10% represents both a pragmatic and conservative choice. For example, a threshold of 0% effect underestimation may discourage further testing, whereas a higher threshold (e.g., 50%) may lead to risk underestimation in the absence of additional model performance criteria. It should be stressed that the methodology as such is independent of the choice of a specific threshold. That can easily be tested within our approach, TWAc-check, on a broader data set if deemed appropriate by regulatory authorities.

The chosen strategy that the 95% confidence interval needs to overlap with the 1:1 line allows for experimental variability to be taken into account. At the same time, it has the disadvantage that data sets with higher variability will more often pass the acceptability criteria (e.g., effects of phenmedipham on *D. magna*; Figure 5A). That is, if there is high variability in the measured data, wider 95% confidence intervals are achieved. Therefore, another criterion may need to be defined to overcome this issue. For example, regardless of the 95% confidence interval, the difference between the measured effect (%) and the predicted effect (%) should be smaller than 25% if the points lie above the 1:1 line. However, this criterion must be finally defined and decided on by the stakeholders involved. Therefore, it has not yet been considered in the examples.

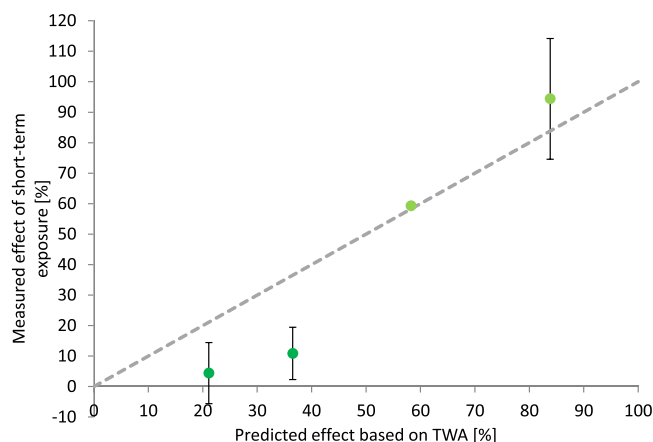


FIGURE 6: Predicted–measured diagram of juvenile survival of *Danio rerio* for spiroxamine based on a 7-day time-weighted average (TWA) concentration.

Following these results, it should also be ensured (for risk assessment applications) that the peak exposure studies are designed such that they contain enough information to robustly test the predictive power of the TWAc. This means that studies should aim not only for low and high effects, but also for at least one effect level in between. However, a higher number of intermediate effect levels would be preferred.

Because the question of applicability of TWAc should be answered on a case-by-case basis (compound and species specific) and not be based on extrapolation between toxicants, scenarios, or even endpoints; we suggest using TWAc-check as a straightforward tool to justify the use of the TWAc in risk assessment. Different pulse exposure scenarios, for example, with intermittent pulse exposures, different peak exposure intervals, or other mechanisms of toxicity, can be tested and used with TWAc-check.

If TWAc-check determines that application of TWAc is not appropriate, different ways forward are possible. We proposed and applied TWAc-check with the default 7-day time window length, as suggested by the EFSA AGD (EFSA, 2013), which also provides the option of adjusting the length of the default TWA period “when scientific data are made available that demonstrate that another TWA period is more appropriate.” Building on this option, TWAc-check can be used to verify whether a shorter TWA window is more appropriate in a specific situation.

If a shorter TWA window does not pass TWAc-check, it can be concluded that the mechanisms involved cannot be appropriately or conservatively described by the simple TWAc model. More sophisticated (Tier 2) tools such as toxicokinetic–toxicodynamic (TKTD) models (Dohmen et al., 2015; EFSA, 2018) could then be used to predict the effects from the refined exposure experiments. The TKTD models can link the external concentration via TK and TD modules to the observed effect (Ashauer et al., 2011, 2016; Brock et al., 2010; Ducrot et al., 2015; Jager et al., 2011; Zimmer et al., 2018). The external concentration is translated via the TK module to the internal concentration, which is the crucial concentration that leads to the observed effects. The TK modules can have different levels of complexity (one-compartment, multicompartment, physiologically based) depending on the level of detail necessary for the research question. Via this module it is possible to assess whether repeated pulses are dependent on or independent from each other, resulting in different effect patterns in the experiment. The TD module depicts the time course of the effect itself. Thus effects might also be observed after transfer to clean medium or full elimination of the toxicant from the water phase. These models are finding wide application in answering scientific questions connected with pulse exposure pattern or fluctuating concentrations in natural habitats (Ashauer et al., 2020), but also increasingly in regulatory contexts. Therefore effect models can also be used (Tier 2) to address the question of TWAc in a regulatory context (see Schmitt et al., 2013 for macrophytes), but with a much higher degree of complexity. Already, the integration of a simple and straightforward method such as TWAc-check can aid in decision-making processes and can act as a complementary tool for decision-making at Tier 1 for the

assessment of all nontarget organisms before moving to the use of more complex models.

CONCLUSIONS

A new approach, TWAc-check, for checking the applicability of TWAc in chronic risk assessment was presented. It directly assesses the degree of conservatism of TWAc using a simple and straightforward approach. With real-example applications, predictions regarding the protectiveness of the TWAc revealed that using TWAc is a case-by-case decision. We demonstrated that TWAc-check can help to make this decision by directly inferring the conservativeness of the TWAc.

Supporting Information—The Supporting Information is available on the Wiley Online Library at <https://doi.org/10.1002/etc.5346>.

Acknowledgment—The authors thank V. Ducrot and L. Lagadic for their helpful scientific contributions, Z. Gao for statistical advice and A. Bone for proofreading the manuscript. Financial support was provided by Bayer CropScience.

Disclaimer—The authors declare the following financial interests, which may be considered as potential competing interests: The project was conducted by Bayer CropScience division. Some of the authors are employed at Bayer or were paid for their work by Bayer.

Author Contributions Statement—**Faten Gabsi**: Conceptualization; Investigation; Data curation; Writing—original draft. **Thomas G. Preuss**: Conceptualization; Investigation; Data curation; Writing—review & editing. **Yvonne Wolf**: Investigation; Data curation; Writing—original draft. **Eric Bruns**: Investigation; Data curation; Writing—review & editing. **Simon Heine**: Investigation; Data curation; Writing—review & editing. **Andreas Solga**: Investigation; Data curation; Writing—review & editing. **Johannes Witt**: Investigation; Data curation; Writing—review & editing.

Data Availability Statement—Data, associated metadata, and calculation tools are available from the corresponding author (yvonne.wolf@bayer.com). The full study reports can be requested for noncommercial use via the Bayer CropScience transparency initiative (email: crops-science-transparency@bayer.com) by referencing the present study and citing the respective M-Number of the study. This registration step is necessary to prevent commercial use of the studies by competitors of the study owner.

REFERENCES

- Ashauer, R., Agatz, A., Albert, C., Ducrot, V., Galic, N., Hendriks, J., Jager, T., Kretschmann, A., O'Conner, I., Rubach, M. N., Nyman, A.-M., Schmitt, W., Stadnicka, J., Van den Brink, P. J., & Preuss, T. G. (2011). Toxicokinetic-toxicodynamic modelling of quantal and graded sublethal endpoints: A brief discussion of concepts. *Environmental Toxicology and Chemistry*, 30(11), 2519–2524. <https://doi.org/10.1002/etc.639>
- Ashauer, R., Albert, C., Augustine, S., Cedergreen, N., Charles, S., Ducrot, V., Focks, A., Gabsi, F., Gergs, A., Goussen, B., Jager, T., Kramer, N. I., Nyman, A.-M., Poulsen, V., Reichenberger, S., Schäfer, R. B., Van den Brink, P. J., Veltman, K., Vogel, S., Zimmer, E. I., & Preuss, T. G. (2016). Modelling survival: Exposure pattern, species sensitivity and uncertainty. *Scientific Reports*, 6, 29178. <https://doi.org/10.1038/srep29178>
- Ashauer, R., Kuhl, R., Zimmer, E., & Junghans, M. (2020). Effect modelling quantifies the difference between the toxicity of average pesticide concentrations and time-variable exposures from water quality monitoring. *Environmental Toxicology and Chemistry*, 39(11), 2158–2168. <https://doi.org/10.1002/etc.4838>
- Bayer. (2006). Toxicity of BYH 18636 technical to duckweed (*Lemna gibba* G3) under static-renewal conditions. Unpublished Report M-269681-01-1.
- Bayer. (2009). Zebra fish (*Danio rerio*), life cycle test, flow through conditions. Unpublished Report M-304458-02-1.
- Bayer. (2013a). *Lemna gibba* G3—Prolonged growth inhibition test with foramsulfuron (AE F130360) with stepwise decreasing concentrations over a 6 week test duration. Unpublished Report M-464150-01-1.
- Bayer. (2013b). *Lemna gibba* G3—Growth inhibition test with foramsulfuron (tech) (AE F 130360) under peak exposure conditions. Unpublished Report M-462569-01-1.
- Bayer. (2013c). *Lemna gibba* G3—Prolonged growth inhibition test with iodosulfuron-methyl-sodium (AE F115008) with stepwise decreasing concentrations and metsulfuron-methyl (AE F075736) with stepwise increasing concentrations over a 6 week test duration. Unpublished Report M-469584-02-1.
- Bayer. (2013d). *Lemna gibba* G3—Prolonged growth inhibition test with mesosulfuron-methyl AE F130060 with stepwise decreasing concentrations over an 8 week test. Unpublished Report M-445139-01-1.
- Bayer. (2014a). Chronic toxicity of phenmedipham to *Daphnia magna* under flow-through conditions. Unpublished report M-482048-01-1.
- Bayer. (2014b). Zebrafish (*Danio rerio*), full life cycle test under static conditions in a water sediment system—Test item: Spiroamine. Unpublished Report M-467979-03-1.
- Bayer. (2016a). *Lemna gibba* G3—Growth inhibition test with thiencarbazone-methyl tech. (BCS-AG17468) under peak exposure conditions—Final report. Unpublished Report M-568404-01-1.
- Bayer. (2016b). *Lemna gibba* G3—Growth inhibition test with iodosulfuron-methyl-sodium TC (BCS-BB66887) under peak exposure conditions. Unpublished Report M-574865-01-1.
- Bayer. (2016c). *Lemna gibba* G3—Growth inhibition test with mesosulfuron-methyl tech. (BCS-AK65185) under peak exposure conditions. Unpublished Report M-577164-01-1.
- Bayer. (2016d). Effects of phenmedipham (techn.) on development and reproductive output of the waterflea *Daphnia magna* under peak-exposure conditions. Unpublished report M-482048-01-1.
- Bayer. (2016e). Effects of aclonifen tech. (BCS-AG74518) on development and reproductive output of the waterflea *Daphnia magna* under continuous static-renewal exposure and under peak exposure conditions. Unpublished report M-566035-01-1.
- Bayer. (2017). *Lemna gibba* G3—Growth inhibition test with foramsulfuron tech. (BCS-AH47624) under peak exposure conditions. Unpublished Report M-572386-03-1.
- Bayer. (2019a). Aclonifen: Toxicity to the water flea *Daphnia magna* Straus under laboratory conditions (21 d reproduction test—Pulsed exposure-scenario). Unpublished report M-670399-01-1.
- Bayer. (2019b). Aclonifen: Toxicity to the water flea *Daphnia magna* straus under laboratory conditions (21 d reproduction test—Pulsed exposure-scenario). Unpublished report M-670403-01-1.
- Belgers, J. D. M., Aalderink, G. H., Arts, G. H. P., & Brock, T. C. M. (2011). Can time-weighted average concentrations be used to assess the risks of metsulfuron-methyl to *Myriophyllum spicatum* under different time-variable exposure regimes? *Chemosphere*, 85(6), 1017–1025. <https://doi.org/10.1016/j.chemosphere.2011.07.025>
- Bliss, C. I. (1934). The method of probits. *Science*, 79(2037), 38–39.
- Brock, T. C. M., Alix, A., Brown, C. D., Capri, E., Gottesbueren, B. F. F., Heimbach, F., Lythgo, C. M., Schulz, R., & Streløke, M. (2010). Linking aquatic exposure and effects—Risk assessment of pesticides. *European Union workshop on linking aquatic exposure and effects in the*

- registration procedure of plant protection products (ELINK). SETAC Press & CRC Press, Taylor & Francis Group.
- Dohmen, G. P., Preuss, T. G., Hamer, M., Galic, N., Strauss, T., Van den Bring, P. J., De Laender, F., & Bopp, S. (2015). Population-level effects and recovery of aquatic invertebrates after multiple applications of an insecticide. *Integrated Environmental Assessment and Management*, 12(1), 67–81. <https://doi.org/10.1002/ieam.1676>
- Ducrot, V., Ashauer, R., Bednarska, A. J., Hinarejos, S., Thorbek, P., & Weymann, G. (2015). Using toxicokinetic-toxicodynamic modeling as an acute risk assessment refinement approach in vertebrate ecological risk assessment. *Integrated Environmental Assessment and Management*, 12(1), 32–45. <https://doi.org/10.1002/ieam.1641>
- European Food Safety Authority. (2013). Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Panel on Plant Protection Products and their Residues (PPR). *EFSA Journal*, 11(7), 3290. <https://doi.org/10.2903/j.efsa.2013.3290>
- European Food Safety Authority. (2015). *Technical report on the outcome of the pesticides peer review meeting on general recurring issues in ecotoxicology*. EFSA supporting publication 2015: EN-924.
- European Food Safety Authority. (2018). Scientific opinion on the state of the art of toxicokinetic/toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. *EFSA Journal*, 16(8), 5377. <https://doi.org/10.2903/j.efsa.2018.5377>
- Jager, T., Albert, C., Preuss, T. G., & Ashauer, R. (2011). General unified threshold model of survival—A toxicokinetic-toxicodynamic framework for ecotoxicology. *Environmental Science & Technology*, 45(7), 2529–2540. <https://doi.org/10.1021/es103092a>
- Organisation for Economic Co-operation and Development. (2002). Test No. 221: *Lemna* sp. growth inhibition test. *OECD guidelines for the testing of chemicals*, Section 2.
- Organisation for Economic Co-operation and Development. (2012). Test No. 211: *Daphnia magna* reproduction test. *OECD guidelines for the testing of chemicals*, Section 2.
- Schmitt, W., Bruns, E., Dollinger, M., & Sowig, P. (2013). Mechanistic TK/TD-model simulating the effect of growth inhibitors on *Lemna* populations. *Ecological Modelling*, 255, 1–10. <https://doi.org/10.1016/j.ecolmodel.2013.01.017>
- Valenzuela, E. F., De Paula, F. F., Teixeira, A. P. C., Menezes, H. C., & Cardeal, Z. L. (2021). Assessment of pesticides in water using time-weighted average calibration of passive sampling devices manufactured with carbon nanomaterial coating on stainless steel wire. *Analytical and Bioanalytical Chemistry*, 413, 3315–3327. <https://doi.org/10.1007/s00216-021-03270-2>
- Watanabe, H., Oda, S., Abe, R., Tanaka, Y., & Tatarazako, N. (2018). Comparison of the effects of constant and pulsed exposure with equivalent time-weighted average concentrations of the juvenile hormone analogue pyriproxyfen on the reproduction of *Daphnia magna*. *Chemosphere*, 195, 810–816. <https://doi.org/10.1016/j.chemosphere.2017.12.124>
- Zimmer, E. I., Preuss, T. G., Norman, S., Minten, B., & Ducrot, V. (2018). Modelling effects of time-variable exposure to the pyrethroid beta-cyfluthrin on rainbow trout early life stages. *Environmental Science Europe*, 30, 36. <https://doi.org/10.1186/s12302-018-0162-0>